

N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM
MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT
CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED
IN THE INTEREST OF MAKING AVAILABLE AS MUCH
INFORMATION AS POSSIBLE

"Made available under NASA sponsorship
in the interest of early and wide dis-
semination of Earth Resources Survey
Program information and without liability
for any use made thereof."

80-16245

DESIGN SPECIFICATION

NASA CR-

FCR

160706

COLOR CODED SPECTRAL PLOTS

Job Order 71-695

N80-30828

(E80-10245) DESIGN SPECIFICATION FOR COLOR
CODED SPECTRAL PLOTS (Lockheed Electronics
Co.) 28 p HC A03/MF A01 CSCL 08B

Unclass
00245

G3/43

Prepared By
Lockheed Electronics Company, Inc.
Aerospace Systems Division
Houston, Texas
Contract NAS 9-12200
For
EARTH OBSERVATIONS DIVISION



National Aeronautics and Space Administration
LYNDON B. JOHNSON SPACE CENTER
Houston, Texas

September 1976

LEC-9924

JSC-11851

DESIGN SPECIFICATION
FOR
COLOR CODED SPECTRAL PLOTS

Job Order 71-695

(AD 63-0997-1695-09)

PREPARED BY

B. E. Wills
B. E. Wills

APPROVED BY

P. L. Krumm
P. L. Krumm, Supervisor
Software Development Section

B. L. Carroll
B. L. Carroll, Manager
LACIE Development and Eval. Dept.

Prepared By
Lockheed Electronics Company, Inc.
For
Earth Observations Division

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

September 1976

LEC- 9924

CONTENTS

Section	Page
1. SCOPE	1-1
2. APPLICABLE DOCUMENTS	2-1
3. SYSTEM DESCRIPTION.	3-1
3.1 <u>HARDWARE DESCRIPTION</u>	3-1
3.2 <u>SOFTWARE DESCRIPTION</u>	3-1
3.2.1 SOFTWARE COMPONENT NO. 1 (NDHIST) .	3-2
3.2.1.1 <u>Linkages</u>	3-2
3.2.1.2 <u>Interfaces</u>	3-2
3.2.1.3 <u>Inputs</u>	3-2
3.2.1.4 <u>Outputs</u>	3-6
3.2.1.5 <u>Storage Requirements</u>	3-8
3.2.1.6 <u>Description</u>	3-9
3.2.1.7 <u>Flowchart</u>	3-10
3.2.2 SOFTWARE COMPONENT NO. 2 (S\TRPL) .	3-11
3.2.2.1 <u>Linkages</u>	3-11
3.2.2.2 <u>Interfaces</u>	3-12
3.2.2.3 <u>Inputs</u>	3-12
3.2.2.4 <u>Outputs</u>	3-16
3.2.2.5 <u>Storage Requirements</u>	3-16
3.2.2.6 <u>Description</u>	3-16
3.2.2.7 <u>Flowchart</u>	3-18
3.2.2.8 <u>Listing</u>	3-19

1. SCOPE

This specification establishes the preliminary design for two axis color coded spectral plots to be output on a multi-file Universal formatted tape.

The location of each pixel on the two axis scatter plot will be computed using the radiance values or two linear combinations of radiance values. The color of the pixel on the scatter plot may optionally be set in one of the following ways:

1. The original radiance value of the pixel.
2. The mean value of the cluster/subclass that the pixel was assigned to during clustering/classification.
3. The mean value of the field the pixel was extracted from for training/test fields.
4. User defined colors.
5. From a different pass when using multi-registered Landsat data.

Offset factors and scale factors may be used for centering the data on the scatter plot. The background may be set to black or may default to white.

The channels will be output on the tape as follows:

1. Channels 1 thru N-1 will be the color of the pixel.
2. Channel N will be the number of occurrences of the pixel.

A color key will be provided when the color is assigned to a pixel by the field mean, cluster mean, or classification re-

sults.

The secondary function is to retrieve information relating to the color codes from the cluster/subclass/field number, image data tape, or mean statistics for the test/training fields.

If cluster/subclass number is requested, a cluster/classification map tape must be input. The map tape is read into core and stored on high speed drum. When a new data vector is encountered during the histogramming processor, the cluster/subclass number for that pixel is retrieved from the drum.

If the field number is requested, the program supplies the field number. If the mean statistics for the fields are requested, the program computes the means of each input field and supplies the field number.

If the color codes are to be extracted from the imagery data tape, the user may select a maximum of four channels. From 1 to 4 radiance values will be packed into one computer word.

For the class level grouping, a file is written each time all the fields within a class are processed; for the subclass level grouping, each time all the fields for one subclass are processed; and for the per field level, each time a field is processed.

A line printer pixel frequency scatter plot may be selected. The location of the pixel is determined by its radiance values the frequency of occurrence or log of frequency of occurrence will be represented by increasing grey levels. A key relating

the grey levels to the number of occurrences of a pixel will be printed.

The above specifications will be implemented into EOD-LARSYS by authorization of Action Document 63-0997-1695-09 in support of Research Test and Evaluation in Building 17/JSC.

This document assumes the reader is familiar with the EOD-LARSYS system and the terminology used in the pattern recognition system.

1-3
3

2. APPLICABLE DOCUMENTS

The following documents, of exact issue shown, form a part of this specification to the extent specified herein:

- User Documentation EOD-LARsys, Lockheed Electronics Company, Inc., HASD, Houston, Texas, November 1975
- Action Document 63-0997-1695-09

3. SYSTEM DESCRIPTION

3.1 HARDWARE DESCRIPTION

N/A

3.2 SOFTWARE DESCRIPTION

The batch oriented, pattern recognition system EOD-LARSHYS is operational on the UNIVAC 1108, Executive II version. At present, the system is composed of nine processors. A processor is defined as an executable program that performs one main function.

To implement the requirements outlined in section 1, two new processors will be added to the existing system. One processor, NDHIST, will perform an N-dimensional histogram of the imagery data ($n \leq 16$) and output a file to be used as an interface to the processor, SCTRPL. The other processor, SCTRPL, will read the N-dimensional histogram file and output two axis color coded spectral plots on a multi-file Universal formatted tape.

3-1
5

3.2.1 SOFTWARE COMPONENT NO. 1 (NDHIST)

The NDHIST processor will perform an N-dimensional histogram of data vectors extracted from the MSS data tape. There are two sets of channels, S_1 and S_2 , used in the computations of histogramming the data. The procedure is as follows:

1. Given $S_1 = \{n_1, n_2, \dots, n_i\}$ and $S_2 = \{\}$, the frequency of occurrence for each unique data vector is computed for channels n_1, n_2, \dots, n_i .
2. Given $S_1 = \{n_1, n_2, \dots, n_i\}$ and $S_2 = \{m_1, m_2, \dots, m_j\}$, the number of unique data vectors accumulated is determined by channels n_1, n_2, \dots, n_i , and the frequency of occurrence associated with each unique data vector is computed using channels m_1, m_2, \dots, m_j . Given x_k and y_k are data vectors determined by sets S_1 and S_2 respectively, the frequency is calculated in one of three ways:
 - a. If for each x_k , there is only one y_k , the frequency is calculated on each occurrence of y_k .
 - b. If for a given x_k , there is more than one y_k , the frequency is calculated on the first y_k encountered. The other y_k 's are ignored.

Example:

Given

$x_1 = \{24, 23, 22, 10\}$ $y_1 = \{41, 42, 43, 20\}$
and

$x_1 = \{24, 23, 22, 10\}$ $y_2 = \{41, 40, 42, 19\}$,

the frequency count associated with x_1 is 1; the vector y_2 is thrown out.

- c. If for a given y_k , there are more than one x_k , the frequency for y_k is calculated based on its association with a particular x_k .

Example:

Given

$x_1 = \{24, 24, 25, 12\}$ $y_1 = \{44, 44, 45, 21\}$
and

$x_2 = \{26, 24, 26, 10\}$ $y_1 = \{44, 44, 45, 21\}$,
the frequency count associated with x_1 is 1; the
frequency count associated with x_2 is 1.

The areas to histogram will be defined by test/training fields. The manner in which the fields are collected or grouped for histogramming is user controlled by input parameters. The data vectors may be collectively histogrammed on class, subclass, or per field level. The maximum number of fields input on any level is 200, and the maximum number of unique data vectors accumulated on any level is 12,000/ (1/4 number of channels). (Maximum of four radiance values per computer word are packed together; each radiance value is stored in nine bits).

3.2.1.1 Linkages

The processor NDHIST will use the Fortran-V compiler, Univac software system routines, and EOD-LARSHYS utility routines.

3.2.1.2 Interfaces

A map tape may optionally be input. If created by ISOCLS, the map tape will contain the cluster images; if created by CLASSIFY, the classification images. See section D of Appendices in the EOD-LARSHYS document for the format of the map tape.

3.2.1.3 Inputs

An imagery data tape must be input. The format must be either Universal or LARSHYS II. See section A and B of Appendices in EOD-LARSHYS document for tape formats.

The card inputs consists of three types of cards: processor, control and field cards.

The processor card is as follows with the keyword starting in column 1:

<u>Keyword</u>	<u>Function</u>
\$NDHIST	Loads all the routines needed by this processor into the system.

With the keywords starting in column 1, the parameters/values starting in any column past 10, and m and n being integer values, the format of the control cards is as follows:

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
CHANNEL	PLOT= n_1, n_2, \dots, n_N , COLOR= m_1, m_2, \dots, m_M (Default: For PLOT channels, no default; for COLOR channels, first 4 PLOT channels, when applicable).	n_i are the channels for determining the position of the pixel ($i \leq 16$). M_j are the channels for the color codes ($j \leq 4$). If $m_j \neq \{ \}$, the histogram is calculated using the COLOR channels; if $m_j = \{ \}$, the histogram is calculated using the PLOT channels.
DATAFILE	UNIT=n,FILE=m (Default: n=3 m=1)	n is the logical unit number assigned to the imagery data tape. m is the file number of the data to process.
MAPFILE	UNIT=n,FILE=m (Default: NONE)	n is the logical unit number assigned to the cluster/classification map tape. m is the file number of the data to process. (The order of the fields to be histogrammed must correspond to the order of the clustered/classified fields on the input map tape)

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
HISFILE	UNIT=n (Default: n=4)	n is the logical unit number assigned to the n-dimensional histogram file output by this processor.
OPTION	CLASS (Default: Field bases)	Fields will be histogrammed on class bases.
OPTION	SUBCLS (Default: Field bases)	Fields will be histogrammed on subclass bases.
OPTION	FIELD (Default: Field bases)	Fields will be histogrammed on per field bases.
OPTION	MEANS	The means of the field(s) will be computed and output on the n-dimensional histogram file.
HED1	Any 60 characters (Default: Lyndon B. Johnson Space Center)	First line of the heading on line printer output.
HED2	Any 60 characters (Default: Houston, Texas)	Second line of the heading on line printer output.
DATE	Any 12 characters (Default: Present date)	Date in the heading on line printer output.

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
COMMENT	Any 60 characters (Default: NONE)	Comment printed with heading on line printer output.
END		Indicates the end of the control card inputs.
\$END*		Indicates the end of all the card inputs.

The field card(s) define the area(s) to be histogrammed and the OPTION control card determines the level of histogramming. The fields may be ordered in one of four ways:

- a. As input to STAT
- b. As input to ISOCLS
- c. As input to CLASSIFY
- d. As test/training fields

Example:

```

CLASS NAME      WHEAT
SUBCLASS       WHEAT1
               field card  1
               field card  2
SUBCLASS       WHEAT2
               field card  3
CLASSNAME      NONWHT
SUBCLASS       NONWH1
               field card  4
SUBCLASS       NONWH2
               field card  5
SUBCLASS       NONWH3
               field card  6
               field card  7

```

If the histogram is accumulated on class bases, fields 1, 2 and 3 are histogrammed collectively and output as data file 1; fields 4, 5, 6 and 7 are histogrammed collectively and output as data file 2.

If the histogram is accumulated on subclass bases, fields 1 and 2 are histogrammed collectively and output as data file 1; field 3 is histogrammed and output as data file 2; field 4 is histogrammed and output as data file 3; field 5 is histogrammed and output as data file 4; fields 6 and 7 are histogrammed collectively and output as data file 5.

If the histogram is performed on per field bases, each field is histogrammed separately and output to a file. Making a total of 7 data files created.

On an accumulative histogram, a maximum of 200 fields may be input.

See section 3.1.3 in EOD-LARSH document for format of field card.

3.2.1.4 Outputs

A multi-file tape is output as an interface to the scatter plot processor, SCTRPL. The file(s) are created with an unformatted Fortran write. The number of records per file depends on the program options selected by the user. No file skipping capability is available; the first file created is always file 1. The format of the tape is as follows:

34
11

file 1 HEADER RECORD
 E-O-F

data file 1

{ RECORD 1
RECORD 2
RECORD 3 (optional)
RECORD 4
RECORD 5
RECORD 6
RECORD 7 (optional)
E-O-F
.
.
.
data file N E-O-F

The contents of each record is as follows:

Header Record

TOTMNS	- Total number of means computed
SIZE	- NOFET2/4
NOFET2	- Number of channels to histogram
(FETVC2(I), I=1, NOFET2)	- Actual channels to histogram
NCLRCH	- Number of color code channels
(CLRVEC(I), I=1, NCLRCH)	- Actual color code channels

Record 1

NOFLD2	- Number of fields histogrammed
NOSUB2	- Number of subclasses histogrammed
TOTUT2	- Number of vertices
NOVEC	- Number of unique vectors histogrammed

Record 2

CL SVC2	- Class name
(SUBVC2(I), I=1, NOSUB2)	- Subclass names
((FIELDS(I, J), I=1, 4), J=1, NOFLD2)	- Field information
((VERTEX(I, J), I=1, 2), J=1, TOTVT2)	- Field vertices

Record 3 (optional)

(MEANS(I), I=1, TOTMNS) - Mean stats for input fields

Record 4

((PLOT(I,J), I=1, SIZE), J=1, NOVEC) - Data vectors

Record 5

(ID(I), I=1, NOVEC) - Class/subclass/field the data vectors belong to

Record 6

(COUNTR(I), I=1, NOVEC) - Number of occurrences of the data vectors

Record 7 (optional)

(COLOR(I), I=1, NOVEC) - Color codes extracted from MSS data tape

As line printer output, the following information will be printed:

- a. Summary of program options selected
- b. For each data file, identification information such as class name, subclass name, field name and field vertices for each input field; the number of unique data vectors found; and the mean stats for each input field.

3.2.1.5 Storage Requirements

By the complexity of the LARSHS program structure, several routines reside in core at all times. With the addition of this processor and the resident routines, 53K words of storage is required to execute this processor.

3.2.1.6 Description

With nine bits constituting a byte, up to four radiance values are stored into one Univac computer word. A character string comparison is used in histogramming the data vectors such that

$$x_i = x_j$$

if, and only if

$$\sum_{i=1}^N \text{word}_i = \sum_{j=1}^N \text{word}_j$$

where word contains the packed data, $i \neq j$, and $N \leq 4$

The field mean statistics are computed accumulatively for channel i by the equation

$$\mu_i = [(N-1)*\mu_{i-1} + x_{k,i}] / N$$

where

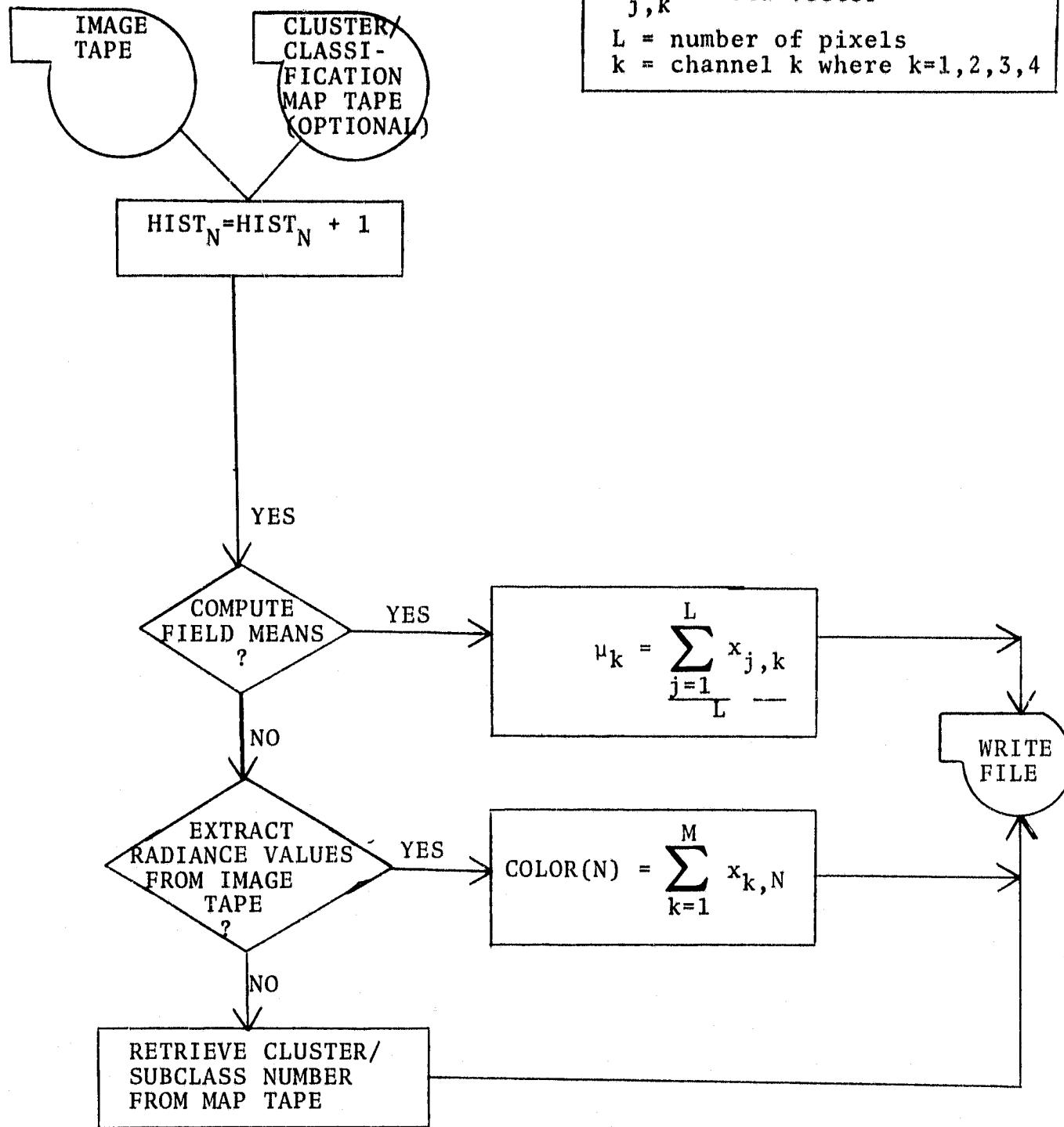
N = number of pixels in the field.

$x_{k,i}$ = radiance value for pixel k, channel i

μ_{i-1} = previously computed mean for channel i

μ_i = new mean for channel i

3.2.1.7 Flowchart



3.2.2 SOFTWARE COMPONENT NO. 2 (SCTRPL)

The SCTRPL processor reads the n-dimensional histogram file written by NDHIST, determines the line bin level and sample bin level for each unique data vector, and outputs the spectral plot in the Universal format. For each file processed on the n-dimensional histogram file, a scatter plot is created and output to tape.

The n-dimensional histogram file is read into core, then stored on high speed drum. Except when transforming the data vectors or sorting the data vectors in descending order, the information from the histogram file is directly accessed from drum as needed.

The data vectors are ordered according to the second coordinate. The second coordinate is the line number for the pixel; the first coordinate, the sample number.

The scatter plot is created line by line. All the pixels belonging to a particular line are collected, and the colors associated with these pixels and number of occurrences of the colors for these pixels are output to tape as channels 1 through N respectively in the sample location determined by the pixel's first coordinate.

A secondary function is to optionally output to the line printer a pixel frequency scatter plot. The frequency of occurrence or log of frequency of occurrence will be represented by a symbol. The location of the symbol on the plot will be determined by the radiance values of the pixel. If the data has been transformed, then the data must be rescale to exercise this option.

3.2.2.1 Linkages

Same as section 3.2.1.1.

3.2.2.2 Interfaces

The n-dimensional histogram file created by NDHIST must be input. See section 3.2.1.4 for format of the tape.

The statistics file created by STAT/ISOCLS may optionally be input. See section 4.1 in EOD-LARSHS document for description of file.

3.2.2.3 Inputs

The card inputs consists of two types of cards: processor and control cards.

The processor card is as follows with the keyword starting in column 1:

<u>Keyword</u>	<u>Function</u>
\$SCTRPL	Loads all the routines needed by this processor into the system.

With the keywords starting in column 1, the parameters/values starting in any column past 10, and m and n being integer values, the format of the control cards is as follows:

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
CHANNEL	$n_1, n_2 \dots n_N$ (Default: First 4 channels from n-dimensional histogram file)	Statistics for these channels will be extracted from the stat file. n_N must be a subset of channels on the stat file.
STATFILE	UNIT=n, FILE=m (Default: NONE)	n is the logical unit number assigned to the stat file. m is the number of the file to process.

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
HISFILE	UNIT=n (Default: n=4)	n is the logical unit number assigned to the n-dimensional histogram file.
PIXPLT	LOG	Line printer pixel scatter plot of the log of frequency of occurrences will be printed.
PIXPLT	FREQ	Line printer pixel scatter plot of the frequency of occurrence will be printed.
PIXPLT	RESCALE (Default: No rescaling. XSIZ=101,YSIZ=101 The range for X-axis is XLO+XSIZ-1; the range for y-axis, YLO+YSIZ-1).	The frequency of occurrences of the pixel for the line printer scatter plot will be scaled to the range XLO, XHI, YLO and YHI. XSIZ will determine the number of bins on the x-axis; YSIZ, the number of bins on the y-axis. (See SIZE control cards).
COLOR	$(m_1), (m_2) \dots (m_N)$ or $L*(m_1), K*(m_{L+1})$ L and K are integer repetition factors.	$m_1 = n_1, n_2 \dots n_M$ - colors assigned to channels 1, 2...M for cluster 1. $m_2 = n_1, n_2 \dots n_M$ - colors assigned to channels 1, 2...M for cluster 2. : $m_N = n_1, n_2 \dots n_M$ - colors assigned to channels 1, 2...M for cluster N. $N \leq 60$ and $M \leq 4$

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
SIZE	XSIZ=n (Default: XSIZ=101)	The number of samples per line to output on the scatter plot tape. $n \leq 200$
SIZE	YSIZ=n (Default: YSIZ=101)	The number of lines to output on the scatter plot tape. $n \leq 200$
SIZE	XHI=n (Default: XHI=100)	The upper limit of the radiance value for the sample axis (x-axis) $n \leq 255$
SIZE	XLO=n (Default: XLO=0)	The lower limit of the radiance value for the sample axis (x-axis) $0 \leq n \leq XHI$.
SIZE	YHI=n (Default: YHI=100)	The upper limit of the radiance value for the line axis (y-axis) $n \leq 255$
	YLO=n (Default: YLO=0)	The lower limit of the radiance value for the line axis (y-axis). $0 \leq n \leq 255$
PLOTAP	UNIT=n (Default: n=7)	n is the logical unit number assigned to the spectral plot tape.

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
B-MATRIX	CARDS (Default: NONE)	CARDS - B-Matrix is being input by cards.
B-MATRIX	FILE	FILE - B-Matrix is being input by file.
BVEC	$t_1, t_2 \dots t_N$	Elements of the additive vector to be used in the transformation. t_i are floating pt. numbers.
BCKGND	n (Default: n=255)	n = 0 - background will be black. n = 255 - background will be white.
SCALE	FILE	The scale factors will be computed from the n-dimensional histogram file.
SCALE	XMAX=t (Default: XMAX will be computed from the n-dimensional histogram file).	The upper range for the transformation of the sample values (x-axis). t is a floating pt. number. See note on next page.
SCALE	XMIN=t (Default: same as above)	The lower range for the transformation of the sample values (x-axis). t is a floating pt. number.

<u>Keyword</u>	<u>Parameter</u>	<u>Function</u>
SCALE	XMIN=t (Default: same as above)	The lower range for the transformation of the sample values (x-axis). t is a floating pt. number.
SCALE	YMAX=t (Default: same as above)	The upper range for the transformation of the line values (y-axis). t is a floating pt. number.
SCALE	YMIN=t (Default: same as above)	The lower range for the transformation of the line values (y-axis). t is a floating pt. number.
SCALE	RESCALE (Default: No rescaling of the transformed data)	The transformed data will be rescaled to the range of XLO, XHI, YLO and YHI. (See SIZE control card)
DATE		
COMMENT		See Section 3.2.1.3.
HED1		
HED2		

Note: If one of the parameters XMIN, XMAX, YMIN or YMAX is input, then all four parameters must be input.

3.2.2.4 Outputs

A multi-file Universal formatted tape containing the spectral plots and color key, when applicable, will be output. The spectral plot images will be the first N lines, followed by M number of blank lines, followed by the color key. See section B of Appendices in EOD-LARSH document for format of tape.

As line printer output the following information will be printed:

- a. Summary of the options selected.
- b. For each file output to tape, identification information such as class name, subclass name, and field information relating to the spectral plot.
- c. Optionally, a line printer pixel frequency scatter plot.

3.2.2.5 Storage Requirements

Same as section 3.2.1.5.

3.2.2.6 Description

If the data vector is to be transformed or reduced to two linear combinations of channels, the following equations and conditions will be applied:

$$y = Bx + c$$

where y is a 2×1 vector
 B is a $2 \times n$ vector, $n \leq 16$
 x is a $n \times 1$ vector, $n \leq 16$
 c is a 2×1 vector

If the transformed data is to be rescaled, the following equation will be applied:

$$y_i = [(HI_i - LO_i) / RANGE_i] * |MIN_i - z_i|$$

where HI_i = an input parameter for the upper limit of the bin level for channel i.

LO_i = an input parameter for the lower limit of the bin level for channel i.

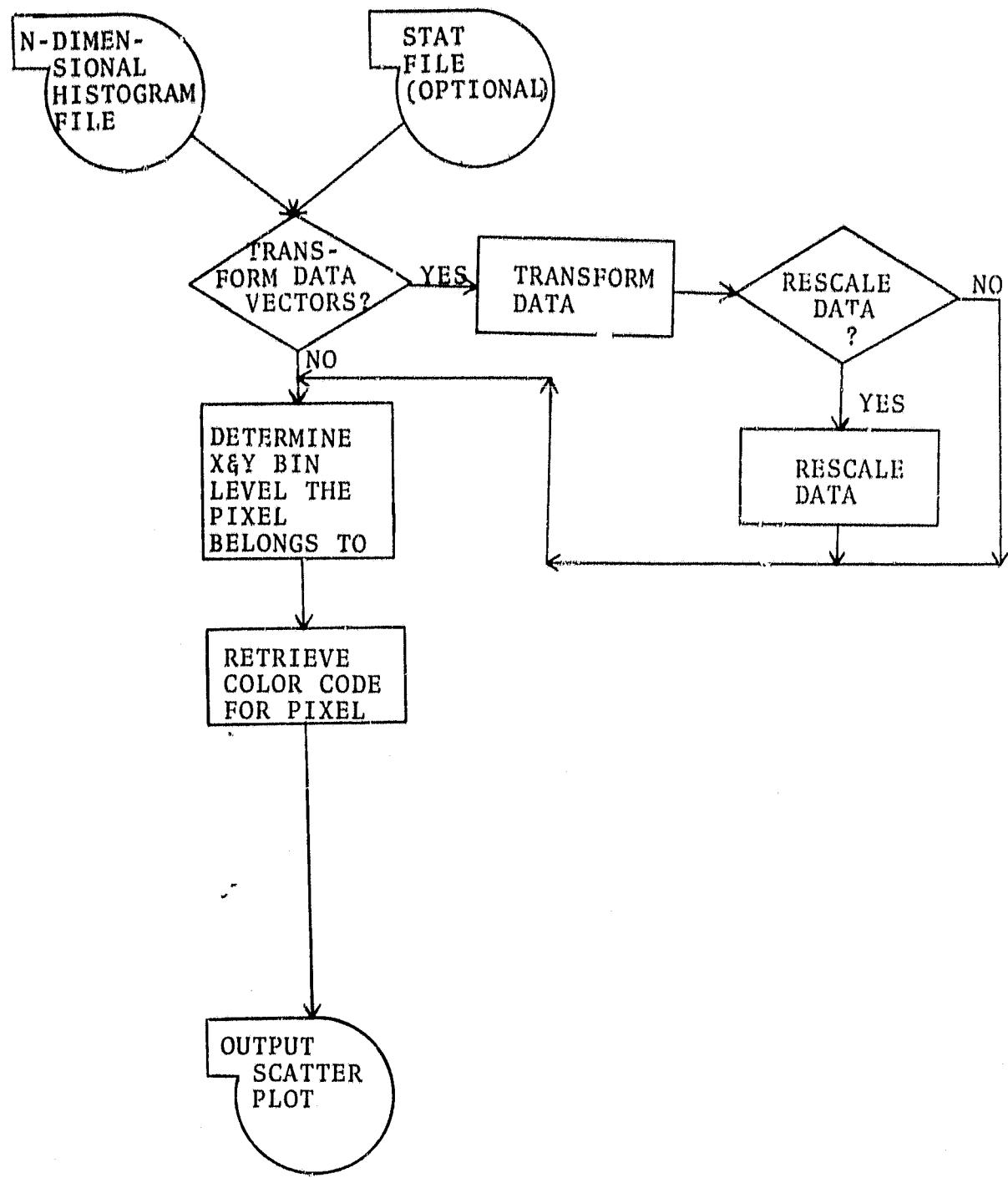
MIN_i = minimum value for channels i.

z_i = transformed data point for channel i.

$RANGE_i$ = range ($\max_i - \min_i$) for channel i.

y_i = rescaled data value for channel i.

3.2.2.7 Flowchart



3.2.2.8 Listing

N/A

3-21
25